

Elżbieta Sobczak

Uniwersytet Ekonomiczny we Wrocławiu

SEGMENTACJA I PROFILOWANIE REGIONÓW EUROPEJSKICH Z WYKORZYSTANIEM METOD KLASYFIKACJI

Streszczenie: Celem opracowania jest segmentacja regionów UE, bazująca na koncepcji analizy portfelowej z wykorzystaniem metod klasyfikacji i analizy dyskryminacyjnej. Podstawę badań stanowią wyniki segmentacji regionalnej ze względu na poziom i dynamikę rozwoju gospodarczego, bazujące na podziale krajów UE na jednostki administracyjne NUTS-2. Zestaw kandydatek na zmienne profilowe obejmuje wybrane wskaźniki konkurencyjności.

1. Wstęp

Macierz rozwoju i udziału w rynku (BCG) opracowana przez Boston Consulting Group jest jedną z powszechniej stosowanych metod oceny pozycji konkurencyjnej przedsiębiorstwa. Konstruowana jest na podstawie dwóch wskaźników: tempa rozwoju rynku i udziału w rynku [McDonald, Dunbar 2003, s. 238; Stonehouse i in. 2001, s. 48-50]. Proponuje się wykorzystanie techniki analizy portfelowej w badaniach segmentacyjnych do wyodrębnienia grup regionów według ich potencjału ekonomicznego oraz dynamiki jego zmian. To podejście umożliwi podział regionów poddawanych analizie na segmenty nadkonkurencyjne, stabilne, „z szansą na rozwój” i mikrozwójowe. Celem opracowania jest segmentacja wybranych regionów Unii Europejskiej, bazująca na koncepcji analizy portfelowej z wykorzystaniem metod klasyfikacji i dyskryminacji.

2. Podstawy metodologiczne regionalnej segmentacji portfelowej

Każdy region poddany segmentacji opisany jest za pomocą następujących zmiennych.

1) $Y = \{Y_1, Y_2\}$ – kryteria segmentacji, Y_1 – udział w otoczeniu, Y_2 – tempo rozwoju.

Tempo rozwoju (Y_2) ustalane jest jako procentowy przyrost względny badanego zjawiska w danym okresie w relacji do okresu przyjętego za podstawę porównań. Proponuje się specyfikację okresu podstawowego z uwzględnieniem następu-

jących zasad: przypisanie mniejszego znaczenia informacjom pochodzącym z wcześniejszych okresów z realizacją zasady postarzania informacji; przypisanie większego znaczenia informacjom pochodzącym z okresów, w których występuje istotna korelacja kryteriów segmentacji z większą liczbą potencjalnych zmiennych profilowych. Identyfikację optymalnego okresu podstawowego, z uwzględnieniem wymienionych zasad, umożliwiła poniższy współczynnik:

$$W_{t+1} = l_{t+1}^2 \cdot c_{t+1} = \frac{l_{t+1}^2}{T-1} \sum_{i=1}^t \frac{1}{T-i} \rightarrow \max, \quad (1)$$

gdzie: W_{t+1} – współczynnik optymalizacji doboru okresu podstawowego, $t = 1, 2, \dots, T-1$ – numer okresu podstawowego, T – liczba okresów poddanych badaniu, l_{t+1} – liczba potencjalnych zmiennych profilowych dostępnych, dostatecznie zmiennych i istotnie skorelowanych z kryterium segmentacji Y_1 i kandydatką na kryterium Y_2 dla $t+1$ okresu podstawowego, c_{t+1} – waga harmoniczna dla $t+1$ okresu podstawowego.

2) $X = \{X_1, X_2, \dots, X_p\}$ – zbiór zmiennych profilowych (deskryptorów).

Procedura badawcza segmentacji portfelowej obejmuje następujące etapy.

Etap I. Dobór i formalno-statystyczna weryfikacja kryteriów segmentacji i potencjalnych zmiennych profilowych. Obejmuje dwustopniową dyskryminację kryteriów segmentacji ze względu na dostępność informacyjną i dostateczną zmienność. Przyjęto, że eliminacji podlegają zmienne, dla których poziom brakujących informacji statystycznych przekracza 15% oraz współczynnik zmienności jest mniejszy niż 10%. Zmienne profilowe powinny być dostępne, dostatecznie zmienne i istotnie skorelowane z kryteriami segmentacji.

Etap II. Podział regionów na segmenty. Segmentacja zostanie przeprowadzona z wykorzystaniem metod analizy skupień. Unitaryzacji kryteriów segmentacji dokonano przez odniesienie ich wartości do rozstępu. Do segmentacji regionów zostaną wykorzystane m.in. hierarchiczne metody analizy skupień: metoda najbliższego sąsiada i metoda Warda. Wymagają one określenia odległości między każdą parą obiektów poddawanych segmentacji. Proponuje się zastosowanie kwadratu odległości euklidesowej. Wybór tej miary implikuje spełnienie tzw. kryterium wariancyjnego, na którym opiera się metoda Warda [Walesiak 1996, s. 110]. Występowanie obiektów nietypowych może zniekształcić wyniki segmentacji, dlatego przed przystąpieniem do klasyfikacji zaleca się ich identyfikację. W tym celu przeprowadzono klasyfikację regionów metodą najbliższego sąsiada, wykazującą tendencję do izolowania obserwacji odstających, przedstawiono wyniki klasyfikacji na dendrogramie oraz wartości kryteriów segmentacji na wykresie punktowym, następnie dokonano merytorycznego wyboru obiektów nietypowych. Określenie liczby segmentów i wstępnych centrów skupień nastąpiło przez:

1) dwukrotną klasyfikację regionów metodą Warda, obejmującą: wszystkie regiony poddawane segmentacji oraz regiony uznane za typowe,

2) przedstawienie wyników klasyfikacji na dendrogramach oraz wykresach odległości wiązania względem etapów wiązania,

3) wstępną wielowariantową propozycję dotyczącą liczby klas i wstępnych centrów skupień, określoną odrębnie dla wszystkich regionów oraz regionów typowych.

Wstępnymi centrami skupień stają się obiekty najbliższe środkom ciężkości klas, do których zostały przyporządkowane. Klasyfikację metodą k -średnich przeprowadzono dla każdego wariantu zaproponowanego w poprzednim kroku procedury badawczej. Do wyboru klasyfikacji optymalnej uzyskanej metodą k -średnich wykorzystano indeks Calińskiego-Harabasa, odrębnie dla wszystkich regionów oraz dla regionów typowych (por. [Caliński, Harabasz 1974, s. 1-27]). Określono stopień podobieństwa otrzymanych klasyfikacji optymalnych z wykorzystaniem skorygowanego indeksu Randa [Hubert, Arabie 1985, s. 193-218]. W przypadku wystąpienia wysokiej zgodności wyników podziałów jako optymalną należy przyjąć klasyfikację wszystkich regionów, zrealizowaną metodą k -średnich, w sytuacji przeciwniej – segmentację otrzymaną z wyodrębnieniem segmentów nietypowych.

Etap III. Analiza składu i struktury otrzymanych segmentów.

Etap IV. Selekcja finalnych zmiennych profilowych z wykorzystaniem wielokrotnej analizy dyskryminacyjnej (por. [Huberty 1994; Krzyśko 1990; Klecka 1980]) i określenie profili segmentów.

3. Segmentacja statyczna wybranych regionów europejskich w 2005 r.

Segmentację przeprowadzono, bazując na regionalnym podziale krajów Unii Europejskiej, dla których punktem odniesienia są jednostki administracyjne NUTS-2. Segmentacja wszystkich regionów europejskich (271) okazała się niemożliwa ze względu na niedostępność informacji statystycznych. W związku z tym segmentacji poddano próbę obejmującą 229 regionów. Przyjęto, że na regiony poddane badaniu składają się jednostki wybrane z niezależnych populacji regionów nadkonkurencyjnych, stabilnych, „z szansą na rozwój” i nierozwojowych. Zrealizowano segmentację statyczną dla 2005 r. Poziom rozwoju regionów reprezentuje kryterium Y_1 – produkt krajowy brutto/1 mieszkańca w jednostkach PPS (UE-27 = 100). Jako zmienne kandydujące do roli kryterium Y_2 reprezentującego tempo rozwoju przyjęto zmienne określające odpowiednio tempo przyrostu produktu krajowego brutto w % (w cenach bieżących) w 2005 r. w relacji do kolejnych lat z okresu 1997-2004. Zaproponowano następujący zestaw kandydatek na zmienne profilowe, obejmujący cechy statystyczne odrębne dla każdego czynnika konkurencyjności: I – czynniki produkcji: X_1 – udział pracujących z wyższym wykształceniem w ogólnej liczbie pracujących w wieku 25-64 lata (w %), X_2 – udział pracujących z wyższym wykształceniem w ogólnej liczbie ludności w wieku 25-64 lata (w %), X_3 – udział ludności dorosłej (w wieku 25-64 lata) uczestniczącej w kształceniu usta-

wicznym w ogólnej liczbie ludności w wieku 25-64 lata (w %), X_4 – zasoby ludzkie w nauce i technologii jako % ogółu ludności, X_5 – zasoby ludzkie w nauce i technologii jako % ludności aktywnej zawodowo w wieku 25-64 lata, X_6 – liczba patentów zgłoszonych do Europejskiego Urzędu Patentowego na 1 mln zasobów siły roboczej; II – klimat społeczno-gospodarczy: X_7 – udział pracujących w usługach związanych z pośrednictwem finansowym, obsługą nieruchomości, wynajmem i działalnością związaną z prowadzeniem interesów w ogóle pracujących (w %), X_8 – stopa bezrobocia (w %), X_9 – bezrobocie długotrwałe (co najmniej 12 miesięcy) jako % bezrobocia ogółem; III – sektory pokrewne i wspomagające: X_{10} – udział pracujących w sektorach wysokich technologii (przemysł i usługach wysokich technologii oraz usługach opartych na wiedzy) w ogóle pracujących (w %), X_{11} – udział pracujących w przemyśle wysokich i średnio wysokich technologii w ogóle pracujących (w %), X_{12} – udział pracujących w usługach opartych na wiedzy i usługach wysokich technologii w ogóle pracujących (w %), X_{13} – udział pracujących w usługach opartych na wiedzy w ogóle pracujących (w %), X_{14} – udział pracujących w usługach rynkowych opartych na wiedzy w ogóle pracujących (w %), X_{15} – udział pracujących w usługach finansowych opartych na wiedzy w ogóle pracujących (w %), X_{16} – wydajność pracy (produkt krajowy brutto na 1 pracującego) w tys. PPS.

Po przeprowadzeniu analizy dostępności informacji statystycznych z potencjalnego zbioru zmiennych profilowych usunięto zmienne X_{10} , X_{12} , X_{15} . Pozostałe kandydatki cechowały się dostępnością informacyjną oraz wystarczającą zmiennością. Wybór optymalnego okresu podstawowego niezbędny dla precyzyjnego określenia kryterium segmentacji Y_2 przeprowadzono na podstawie analizy korelacji potencjalnych zmiennych profilowych z kryterium segmentacji Y_1 i zmiennymi kandydującymi do roli kryterium segmentacji Y_2 . Wartość krytyczną współczynnika korelacji ustalono na poziomie $r_1^* = 0,13$. Jest to wartość graniczna, powyżej której współczynnik korelacji jest istotny statystycznie na poziomie istotności $\alpha = 0,05$, dla $N - 2 = 227$ stopni swobody. Optymalnym okresem podstawowym, o maksymalnej wartości współczynnika W_{t+1} , okazał się rok 2002. Stąd też kryterium segmentacji Y_2 można określić jako tempo przyrostu produktu krajowego brutto w % (w cenach bieżących) w 2005 r. w relacji do roku 2002. Ostatecznie wybrano następujące zmienne profilowe spełniające warunki dostępności, zmienności i istotnej korelacji z kryteriami segmentacji: X_3 , X_4 , X_5 , X_7 , X_{13} , X_{14} , X_{16} .

Dokonano identyfikacji następujących regionów nietypowych: Région de Bruxelles-Capitale, Luksemburg (Grand-Duché), Bratislavský kraj i Inner London. Następnie przeprowadzono klasyfikację 229 i 225 (bez nietypowych) regionów metodą Warda. Zaproponowano następujące wariantowe klasyfikacje: 229 regionów na 7-15 segmentów oraz 225 regionów na 6-10 segmentów. Przeprowadzono je, posługując się metodą k -średnich. Jako wstępne centra skupień ustalono obiekty najbliższe środkom ciężkości grup regionów wyodrębnionych metodą Warda w

danym wariancie segmentacji. Wyboru segmentacji optymalnej dla 229 i 225 regionów dokonano, posługując się indeksem jakości klasyfikacji Calińskiego-Harabasz. Klasyfikacja optymalna 229 regionów UE obejmuje podział na 15 segmentów, a 225 regionów podział na 7 segmentów. Zdecydowano, że obiekty nietypowe potraktowane zostaną jako cztery jednoelementowe segmenty, stąd optymalna segmentacja uwzględniająca regiony nietypowe obejmuje podział na 11 grup. Wartość skorygowanego indeksu Randa wynosi $CR = 0,489$, oznacza niezbyt wysokie podobieństwo porównywanych podziałów. Dlatego ostatecznie jako optymalną przyjęto segmentację obejmującą podział analizowanych regionów na 11 grup (włącznie z czterema segmentami nietypowymi). Wyniki mezosegmentacji optymalnej zestawiono w tab. 1. Na rysunku 1 przedstawiono lokalizację segmentów w macierzy rozwoju i udziału w otoczeniu.

Tabela 1. Segmentacja optymalna regionów Unii Europejskiej w 2005 r.

Segment	Państwo (liczba regionów)	Liczebność (częstość)
1	Włochy (8), Francja (7), Niemcy (5), Belgia (4), W. Brytania (3), Portugalia (3), Malta (1), Grecja (1), Węgry (1), Finlandia (1)	34 (14,8%)
2	Niemcy (6), Austria (4), Niderlandy (4), W. Brytania (4), Belgia (1), Włochy (1), Finlandia (1), Rep. Czeska (1), Irlandia (1), Szwecja (1), Francja (1)	25 (10,9%)
3	W. Brytania (17), Hiszpania (12), Szwecja (5), Niemcy (4), Austria (4), Belgia (1), Węgry (1), Niderlandy (1), Finlandia (1), Portugalia (1), Grecja (1), Francja (1)	49 (21,4%)
4	Rep. Czeska (6), Hiszpania (5), Polska (3), Grecja (2), W. Brytania (1), Cypr (1), Irlandia (1), Węgry (1)	20 (8,7%)
5	Niemcy (19), Włochy (11), Francja (10), Niderlandy (7), W. Brytania (6), Belgia (3), Szwecja (2), Portugalia (1), Grecja (1), Finlandia (1),	61 (26,6%)
6	Rumunia (8), Bułgaria (4), Łotwa (1), Litwa (1), Estonia (1), Rep. Czeska (1), Słowacja (1)	17 (7,4%)
7	Polska (11), Węgry (4), Słowacja (2), Bułgaria (2)	19 (8,3%)
8	Belgia (1)	1 (0,44%)
9	Luksemburg (1)	1 (0,44%)
10	Słowacja (1)	1 (0,44%)
11	W. Brytania (1)	1 (0,44%)
Suma	-	229 (100%)

Źródło: opracowanie własne na podstawie bazy danych Eurostatu.

Najliczniejszymi segmentami są piąty, skupiający 61 regionów, i trzeci – 39 regionów. W segmentach szóstym i siódmym znalazły się wyłącznie regiony z krajów tzw. nowego rozszerzenia Unii Europejskiej. Zdecydowanie najwyższym poziomem PKB *per capita* w relacji do średniej unijnej w 2005 r. cechowały się kolejno jednoelementowe segmenty nadkonkurencyjne, zawierające Inner London

T E M P O	Bardzo wysokie Ponad 20	Segment 6 (17 regionów)	„z szansą na rozwój”	nadkonkurencyjne Segment 10 (1 region)	Segment 9 (1 region)
	Wysokie (13, 20]		Segment 4 (20 regionów)	Segment 3 (49 regionów)	Segment 11 (1 region)
R O Z W O J U	Niskie (8, 13]	Segment 7 (19 regionów)	nierozwojowe	Segment 2 (25 regionów)	stabilne
	Bardzo niskie Do 8		Segment 1 (34 regiony)	Segment 5 (61 regionów)	Segment 8 (1 region)
W %		bardzo niski do 50	niski (50, 100]	wysoki (100, 170]	bardzo wysoki ponad 170
	UDZIAŁ W OTOCZENIU w %				

Rys. 1. Macierz rozwoju i udziału w otoczeniu dla regionów Unii Europejskiej w 2005 r.

Źródło: opracowanie własne.

(302,7%) i Luksemburg (Grand-Duché) (263,3%) oraz stabilny Région de Bruxelles-Capitale (240,5%). Najwyższym tempem przyrostu PKB *per capita* w 2005 r. w porównaniu z rokiem 2002 charakteryzował się nadkonkurencyjny segment dziesiąty (Bratislavský kraj – 33,16%), następnie segment szósty „z szansą na rozwój” (29,74%), zawierający Litwę, Łotwę, Estonię oraz 8 regionów rumuńskich, 4 bułgarskie i po jednym czeskim i słowackim. Bardzo wysokie tempo przyrostu PKB odnotowano również w nadkonkurencyjnym segmencie dziewiątym – Luksemburg (Grand-Duché) (25,53%). Segmenty siódmy i szósty cechowały się najniższym przeciętnym udziałem PKB/1 mieszkańca w średniej unijnej odpowiednio – 41,6 i 42,81%. W segmencie siódmym (nierozwojowym) znajduje się 11 regionów polskich, 4 węgierskie i po 2 słowackie i bułgarskie. Najniższym tempem przyrostu PKB cechował się ósmy segment stabilny (Région de Bruxelles-Capitale – 6,28%), następnie nierozwojowy segment pierwszy (6,66%) i stabilny segment piąty (7,88%).

Do selekcji finalnych zmiennych profilowych zastosowano wieloraką analizę dyskryminacyjną. Przyjęte zmienne profilowe ($X_3, X_4, X_5, X_7, X_{13}, X_{14}, X_{16}$) zostały zmierzzone na skali metrycznej i żadna z nich nie stanowi kombinacji liniowej pozostałych. Poddano weryfikacji kolejne założenia analizy dyskryminacyjnej: zmienne w segmentach powinny wykazać się łącznie wielowymiarowym rozkładem normalnym (skorzystano z testu Shapiro-Wilka) i równością macierzy wariancji i kowariancji (test Boksa i jego przybliżenie rozkładem F Fishera-Snedccora). Założenia te nie zostały spełnione, ponieważ nie są krytyczne, analizę dyskryminacyjną przeprowadzono (z wyłączeniem regionów nietypowych). Zastosowano procedurę krokowej postępującej analizy dyskryminacyjnej. Do modelu wprowadzono wszystkie zmienne profilowe: $X_3, X_4, X_5, X_7, X_{13}, X_{14}, X_{16}$. Oceny istotności zmiennych w dyskrymi-

nowaniu segmentów dokonano na podstawie cząstkowych współczynników Λ_p Wilksa z wykorzystaniem statystyki F_p , mającej rozkład Fishera-Snedecora. Wyniki zestawione w tab. 2 oznaczają, że wszystkie zmienne istotnie wpływają na dyskryminację segmentów. Moc dyskryminacyjna otrzymanego modelu okazała się istotna statystycznie (λ Wilksa $\Lambda = 0,0903$, $\chi^2 = 521,7780$, $\chi^2_{0,05;42} = 58,124$).

Weryfikacji istotności otrzymanych funkcji dyskryminacyjnych dokonano z wykorzystaniem informacji zawartych w tab. 3. Wartości współczynnika udziału poszczególnych funkcji dyskryminacyjnych w ogólnej mocy dyskryminacyjnej (U_z) wskazują, że największe znaczenie ma pierwsza funkcja (92,5%), dużo mniej jest wkład pozostałych funkcji. Do podobnych wniosków prowadzi analiza wartości współczynnika korelacji kanonicznej (R_{cz}). Jedyne trzy pierwsze funkcje dyskryminacyjne wykazują istotność statystyczną, czyli zdolność do separacji wyłonionych segmentów. Pozostałe funkcje takiej własności nie mają ($\chi^2 < \chi^2_{0,05;df}$), dlatego nie będą poddawane interpretacji.

Tabela 2. Podsumowanie analizy funkcji dyskryminacyjnej dla 7 zmiennych profilowych

Zmienne profilowe	Cząstkowy współczynnik Λ_p Wilksa	Statystyka empiryczna F_p	<i>p-value</i>
X_{16}	0,3501	65,5884	0,0000
X_3	0,7284	13,1775	0,0000
X_{13}	0,8160	7,9667	0,0000
X_4	0,8424	6,6122	0,0000
X_5	0,8708	5,2402	0,0000
X_7	0,9291	2,6962	0,0152
X_{14}	0,9313	2,6077	0,0185

Źródło: obliczenia własne z wykorzystaniem pakietu statystycznego STATISTICA 8 PL.

Tabela 3. Weryfikacja istotności funkcji dyskryminacyjnych dla segmentacji regionalnej

Funkcja dyskryminacyjna	R_{cz}	U_z (w %)	Λ Wilksa	χ^2	Stopnie swobody <i>df</i>	$\chi^2_{0,05;df}$	<i>p-value</i>
D_1	0,9251	92,0	0,0903	521,7821	42	58,12	0,0000
D_2	0,4766	4,6	0,6260	101,6361	30	43,77	0,0000
D_3	0,3632	2,4	0,8100	45,7295	20	31,41	0,0009
D_4	0,2070	0,7	0,9331	15,0232	12	21,03	0,2402
D_5	0,1479	0,3	0,9749	5,5224	6	12,59	0,4788
D_6	0,0577	0,1	0,9967	0,7234	2	5,99	0,6965

gdzie: R_{cz} – współczynnik korelacji kanonicznej z-tej funkcji dyskryminacyjnej, U_z – współczynnik udziału z-tej funkcji w ogólnej mocy dyskryminacyjnej wszystkich funkcji.

Źródło: obliczenia własne z wykorzystaniem pakietu statystycznego STATISTICA 8 PL.

Z analizy standaryzowanych współczynników dyskryminacyjnych można wysnuć następujące wnioski. Na pierwszą funkcję, mającą najsilniejszą moc dyskryminacyjną, największy wpływ dodatni wywiera zmienna X_4 , ujemny zmienna X_5 , dodatni zmienna X_{16} , ujemny X_{14} i dodatni X_7 . Na drugą funkcję najsilniejsze oddziaływanie o charakterze ujemnym wywiera zmienna X_{13} , oddziaływanie dodatnie zmienna X_3 . Na trzecią funkcję, podobnie jak na pierwszą, najsilniejszy wpływ dodatni wywiera X_4 , a ujemny X_5 . Podobną analizę przeprowadzono na podstawie współczynników struktury czynnikowej (współczynników korelacji liniowej zmiennych profilowych z wartościami funkcji dyskryminacyjnych). Najsilniejszą korelację dodatnią z pierwszą funkcją dyskryminacyjną wykazały kolejno zmienne X_{16} , X_7 , X_{14} , X_{13} , X_4 , z drugą funkcją zmienne X_3 (wpływ dodatni), X_{16} i X_{13} (wpływ ujemny), z trzecią X_3 i X_{13} . Za finalne zmienne profilowe przyjęto zmienne, których istotny wpływ na daną funkcję dyskryminacyjną wynikał z analizy zarówno współczynników standaryzowanych, jak i współczynników struktury czynnikowej. Należą do nich zmienne X_{16} , X_4 , X_7 dla pierwszej funkcji oraz X_3 i X_{13} dla funkcji drugiej. Profile segmentów zbudowano na podstawie następujących deskryptorów: wydajność pracy na 1 pracującego w tys. PPS (X_{16}), zasoby ludzkie w nauce i technologii jako % ogółu ludności (X_4), udział pracujących w usługach związanych z pośrednictwem finansowym, obsługą nieruchomości, wynajmem i działalnością związaną z prowadzeniem interesów w ogóle pracujących (X_7), udział ludności

Tabela 4. Profile wyodrębnionych segmentów

Zmienne	X_{16}	X_4	X_7	X_3	X_{13}
Segmenty nadkonkurencyjne					
3	52,9	27,2	10,6	14,7	34,8
9	142,3	28,9	18,9	8,5	42,0
10	63,7	36,2	14,7	14,8	39,5
11	155,7	38,9	25,0	24,6	57,4
Segmenty stabilne					
2	71,6	33,4	14,0	13,7	40,1
5	56,5	27,0	10,7	10,1	34,3
8	146,5	38,0	20,0	12,0	48,2
Segmenty „z szansą na rozwój”					
4	39,1	20,4	7,1	6,6	23,7
6	22,7	17,9	4,3	2,7	19,0
Segmenty nirozwojowe					
1	45,9	21,4	7,9	7,4	30,3
7	26,0	17,0	5,7	3,8	23,7

Źródło: obliczenia własne.

dorośle (w wieku 25-64 lata) uczestniczącej w kształceniu ustawicznym w ogólnej liczbie ludności w wieku 25-64 lata (X_3), udział pracujących w usługach opartych na wiedzy w ogóle pracujących (X_{13}). Profile wyodrębnionych segmentów przedstawia tab. 4. Umieszczono w niej również profile segmentów nietypowych (8, 9, 10 i 11), nieuwzględnionych w analizie dyskryminacyjnej.

Regiony nietypowe (8, 9, 10 i 11) powodują bardzo duże zróżnicowanie segmentów nadkonkurencyjnych i stabilnych ze względu na przeciętną wydajność pracy. Nie obserwuje się tego zjawiska w przypadku pozostałych finalnych zmiennych profilowych. Wszystkie zmienne rozróżniają segmenty nadkonkurencyjne i stabilne od tych „z szansą na rozwój” i nierozwojowych, a więc regiony różniące się udziałem w otoczeniu. Natomiast nie separują dobrze segmentów odmiennych ze względu na tempo rozwoju.

4. Podsumowanie

W opracowaniu przedstawiono wyniki segmentacji 229 regionów Unii Europejskiej w 2005 r. Wykorzystano koncepcję analizy portfelowej oraz metody klasyfikacji – metodę k -średnich i analizę dyskryminacyjną. Kryteria segmentacji określone jako PKB/1 mieszkańca i tempo przyrostu PKB pozwoliły na wyodrębnienie 11 segmentów. Najliczniejsza okazała się grupa segmentów stabilnych (87 regionów), następnie nierozwojowych (53) i nadkonkurencyjnych (52). Najmniej liczna jest grupa segmentów „z szansą na rozwój” (37). Zmienne profilowe zdecydowanie lepiej separują segmenty zawierające regiony o niskim i wysokim poziomie PKB/1 mieszkańca niż segmenty regionów o odmiennej dynamice rozwoju.

Literatura

- Caliński R.B., Harabasz J., *A dendrite method for cluster analysis*, „Communications in Statistics” 1974 nr 3, s. 1-27.
- Hubert L., Arabie P., *Comparing partitions*, „Journal of Classification” 1985 nr 2, s. 193-218.
- Huberty C.J., *Applied Discriminant Analysis*, John Wiley & Sons, New York 1994.
- Klecka W.R., *Discriminant Analysis*, Sage Publication, London 1980.
- Krzyśko M., *Analiza dyskryminacyjna*, Wydawnictwo Naukowo-Techniczne, Warszawa 1990.
- McDonald M., Dunbar I., *Segmentacja rynku. Przebieg procesu i wykorzystanie wyników*, Oficyna Ekonomiczna, Kraków 2003.
- Stonehouse G. i in., *Globalizacja. Strategia i zarządzanie*, Felberg SJA, Warszawa 2001.
- Walesiak M., *Metody analizy danych marketingowych*, Wydawnictwo Naukowe PWN, Warszawa 1996.

SEGMENTATION AND PROFILING OF EUROPEAN REGIONS BY MEANS OF IMPLEMENTING CLASSIFICATION METHODS

Summary: Segmentation of the European Union regions based on portfolio analysis concept and implementing discrimination methods for classification and analysis has become the objective of the study. Regional segmentation results with regard to the level and dynamics of economic development, based on the European Union countries division into NUTS 2 administrative units, have become the basis for conducting research. The set of candidates for profile variables includes selected indicators of competitiveness.